Words (text only) 9,083

The Neuronal Basis of Visual Consciousness

Christof Koch[1] and Francis Crick[2]

May 16. 2002

[1]Computation and Neural Systems Program, California Institute of Technology, Pasadena, California 91125 and [2]The Salk Institute, La Jolla, California, 92037.

Corresponding author:
Prof. Christof Koch
Division of Biology, 139-74
Caltech
Pasadena, CA 91125
Tel: (626) 395-6855
FAX: (626) 796-8876
Email: koch@klab.caltech.edu
Web: www.klab.caltech.edu

1. Introduction

We outline in this chapter some of our present ideas about consciousness in general and visual consciousness in particular. For now, we believe that the most productive research strategy is to focus on the neuronal correlates of consciousness (NCC). The next step is to establish the exact nature of the causal relationship between neuronal events and subjective feelings and, finally, to understand the thorny philosophical problem of qualia or subjective feelings and how they can arise out of certain physical systems. We assume that higher mammals, such as rodents or primates, share certain forms of sensory consciousness with humans, even though these animals lack language skills.

To characterize the NCC, we have to contrast neural activity that directly gives rise to conscious sensations, thoughts or action with neural activity that is associated with unconscious, stereotyped and on-line visuo-motor behavior. Where is the difference between these forms? How do these differences in activity relate to the ventral and dorsal streams? We emphasize the importance of explicit representations, the idea of essential nodes in a network and whether such nodes correspond to the columnar properties of a patch of cortex. We also discuss whether the correlated firing of a set of neurons is needed for consciousness and the role of cortical area V1 and prefrontal areas in consciousness. We end by briefly describing some of the relevant experiments. For earlier versions of these ideas, see (Crick and Koch, 1998 and Koch and Crick, 2000).

2. Prolegomenon to any Study of Consciousness

The most puzzling aspect of vision and visual perception is that it frequently gives rise to conscious seeing: to the vivid sense of a hot pink shirt, the sight of the face of a loved one or the profound sense of depth when gingerly stepping up to the edge of a bluff. After a hiatus of half a century or more (e.g. Westheimer, 1999), the physical origins of consciousness are being, once again, vigorously debated innumerous books and monographs published in the last decade. What is it that we can we ascertain about the neurobiological basis of consciousness, in particular visual consciousness, and what can we reasonably assume at this point in time? Neuroscientists have made a number of working assumptions that, in the fullness of time, need to be justified more fully.

Firstly, and most importantly, that there is something to be explained. That the subjective content associated with a conscious sensation exists and has its physical basis in the brain. Philosophers refer to this as *qualia*, *what-it-is-like*, the *aboutness of consciousness* and other terms. At this point in time, no even remotely satisfactorily accounts as to why brains can produce subjective feelings have been made. Our strategy is to focus for now on other aspects of qualia where progress can be made.

There is general agreement that much of consciousness is private. I cannot convey to you exactly how red looks to me, even if experiments show that you and I respond to colors in a very similar way. We have claimed (Crick and Koch, 1995a) that this is because, at each stage of processing in the cortex, the information symbolized is re-coded, so that the more internal neural activity is only expressed very indirectly in any motor output, such as speech. On the other hand a person can say whether two similar shades of red appear identical or not. It is not surprising that much of the content of consciousness is largely private. What is mysterious is the exact nature of these internal experiences.

Secondly, consciousness is one of the principal properties of the human brain, a highly evolved system. It therefore must have some useful function to perform. It is extremely unlikely that the well-structured nature of conscious representations are epiphenomena without any causal function as sometimes asserted by philosophers. We (Crick and Koch, 1995b) assume that the function of visual consciousness is to produce the best current interpretation of the visual scene---in the light of past experiences---and to make it available, for a sufficient time, to the parts of the brain which contemplate, plan and execute voluntary motor outputs (including language). This needs to be contrasted with on-line or zombie behaviors that bypass consciousness but can generate stereotyped motor outputs (see below).

Thirdly, at least some animal species possess some aspects of consciousness. In particular, this is assumed to be true for non-human primates, such as the macaque monkey. Visual consciousness is likely to be similar in humans and monkeys for several reasons. (1) For most sensory tasks (e.g. for visual motion discrimination as discussed in the chapters by **Albright & Dobkins, Britten** and **Orban**) trained monkeys behave as humans do under similar conditions, (2) The gross neuroanatomy of humans and non-human primates is the same once the difference in size has been accounted for. This is not to deny differences in the microcircuitry nor the existence of cell types unique to human and great apes (Nimchinsky *et al.*, 1999). (3) As emphasized throughout this handbook, functional

brain imaging in humans is confirming the existence of a functional organization similar to that characterized by single cell electrophysiology in the monkey.

As a corollary, it follows that neither language nor a highly developed sense of the self or of others is necessary for consciousness to occur. Indeed, even severely autistic individuals can carry out very complex sensory discrimination tasks and show no evidence of deficits in sensory discrimination or detection tasks, conscious or not (Baron-Cohen, 2001). Of course, language does enrich consciousness considerably, but there is little hard evidence that it is necessary for conscious sensations to occur.

3. Conscious and Non-Conscious Visual Behaviors

We can state bluntly the question that neuroscience must ultimately answer: It is probable that at any moment some active neuronal processes in our head correlate with consciousness, while others do not; what is the difference between them? We refer to the minimal neuronal mechanisms necessary and sufficient for any one specific conscious percept as the *neuronal correlates of consciousness* (or the *NCC*). We believe that at this point in time, it is empirically most profitable to seek to identify the NCC and to defer many of the other problems (such as qualia).

It is critical to distinguish levels of consciousness (as in drowsy versus awake) from the NCC associated with a specific phenomenal content. This distinction reflects everyday experience of the distinction between being conscious versus unconscious, on the one hand, and being conscious of a high C versus not being consciousness of this tone. *Enabling factors* are required for any conscious sensation to occur without reflecting any specific one.

Cognitive and clinical research demonstrates that much complex information processing can occur without involving consciousness. In normal subjects, such *unconscious* processing includes priming, many aspects of oculo-motor behaviors, the implicit recognition of complex sequences, posture adjustment, and many highly trained behaviors, such as driving, climbing, dancing, playing tennis and so on (Velmans, 1991; Koch and Crick, 2001).

Milner and Goodale (1995) make a masterful case for the existence of so-called *on-line* actions that by-pass consciousness (see also Rossetti, 1998). Their function is to mediate relative stereotyped visuo-motor behaviors, such as eye and arm movements, posture adjustments and so on, in a very rapid manner. On-line agents---or zombies as we call them---operating in the here and now and lacking direct access to working memory, are complemented by one or more

networks for object classification, recognition and identification. It is these latter ones that give rise to conscious perception.

The exact relationship between zombie and conscious systems is unclear. In one reading, these might coexist as anatomical distinct streams in the brain. Milner and Goodale (1995) suggest that the on-line system mainly uses the dorsal visual stream (see the chapter by **Ungerleider & Pasternak**), implying that all activity in this pathway is unconscious. They consider the ventral stream (see the chapters by **Tanaka** and by **Kanwisher**), on the other hand, to be largely conscious. Another interpretation holds that zombie and conscious systems share the same anatomical substrate but corresponds to different modes. For instance, feedforward activity that rapidly moves through the relevant networks to trigger motor action might well correspond to on-line behaviors while sustained firing activity that is modulated by cortico-cortical feedback, in particular from prefrontal cortex, might be a critical component of the NCC.

The broad properties of the two hypothetical entities -- the on-line and the seeing systems -- are shown in the Table, following Milner and Goodale (1995; for critical evaluation and alternative views, see Boussaoud, di Pellegrino, and Wise, 1996 and Franz *et al.*, 2000). The on-line system may have multiple subcomponents (e.g., for eye movements, for arm movements, for body posture adjustment). Under normal conditions, on-line behaviors coexist with conscious percepts, thoughts or actions. Indeed, there is evidence that in some circumstances the seeing system can interfere with the on-line system (Rossetti, 1998).

	ON-LINE SYSTEM	SEEING SYSTEM
Visual inputs handled	must be simple	can be complex
Motor outputs	stereotyped responses	many possible responses
Minimum time needed for response	short	longer
Effect of a few seconds' delay	may not work	can still work
Coordinates used	egocentric	object-centered
Certain perceptual illusions	not effective	seen

Conscious	no	yes

Comparison of the Hypothetical On-line System and the Seeing System
(based on Milner and Goodale, 1995)

One striking piece of evidence for on-line behaviors comes from studies on patient D.F. Her brain has diffuse damage produced by carbon-monoxide poisoning. She is able to see color and texture very well but is very deficient in seeing orientation and form. In spite of this, she is very good at catching a ball. She can "post" her hand or a card into an elongated slot without difficulty, though she is unable to report the slot's orientation. If the stimulus is removed and DF's is asked to delay her response by a few seconds, she is unable to properly orient her hand. That is, on-line visuo-motor behaviors do not have access to working memory since they operate in the here and now. This observation might be used to distinguish all manner of zombie actions from conscious ones.

We suspect that while these suggestions about two systems are on the right lines, they are probably oversimplified. The little that is known of the neuroanatomy would suggest that there are likely to be multiple cortical streams, with numerous anatomical connections between them (Distler *et al.*, 1993). For example, recent data implicates the superior temporal cortex in humans as critical for spatial awareness, attention and exploration (Karnath, 2001). It is difficult to assign this brain area unambiguously to either the ventral or the dorsal pathway, raising the possibility that it acts as an interface between the two. Furthermore, the neuroanatomy does not suggest that the sole pathway goes up to the highest levels of the visual system, and from there to the highest levels of the prefrontal system and then down to the motor output. There are numerous pathways from intermediate levels of the visual system to intermediate frontal regions.

Other examples of residual visual behaviors in the absence of conscious, visual sensation in brain-damaged patients include *blindsight* (see the chapter by **Weiskrantz**) and the *visual neglect syndrome* (Driver and Mattingley, 1998). In extinction, the presence of a visual stimulus in the intact hemifield successfully competes with the stimulus in the lesioned hemifield: the latter one is simply not perceived. Typically, these symptoms are associated with lesions in the contralateral posterior parietal lobes (for an alternative view, see Karnath, 2001). This has been taken as *prima facie* evidence for the importance of these areas for visual consciousness.

Bilateral damage of both parietal lobes can give rise to *Balint syndrome* or simultanagnosia. Such patients only see one thing at a time, neglecting everything else but the one object that currently captures their attention at the exclusion of anything else (Rafal, 1997). Yet the critical point is that they can still see something. This leads us to conclude that the areas involved in spatial neglect are not strictly necessary for conscious vision *per se*. It is likely, though, that parietal areas provide attentional modulation of processing in the ventral pathway by biasing the competition among neuronal assemblies there.

Finally, two groups of people that appear to function, on occasion, in the absence of conscious sensations are a subset of patients with complex partial seizures (Penfield and Jasper, 1954) and sleep-walkers (Revonuso *et al.*, 2000). Both can execute routine behaviors, including mumbling, walking, moving furniture, driving and other automatic behaviors without being able to interact (e.g. have a conversation) with their environment nor recall anything that occurred during the seizure or the somnambulistic episode. It is possible that the NCC is shut down yet enough specific nervous activity remains to support zombie behaviors (Koch and Crick, 2001).

All of this evidence suggests that many behaviors occur in the absence of conscious sensation. This offers the possibility of a contrastive program, in which the neural mechanisms underlying automatic or zombie-behaviors are contrasted with those underlying any one conscious sensation, the NCC proper. However, the existence of unconscious behaviors also implies that one has to exercise great caution when working with animals, babies or severely aphasic patients. It therefore becomes important to develop criteria to distinguish opto-motor behaviors controlled by zombie agents from those that require consciousness. Interposing a delay of a few seconds between the visual stimulus and the required response seems currently the most promising litmus test.

The existence of a host of zombie behaviors also raises the question of why we need consciousness at all? Why aren't we large bundles of on-line behaviors? We proposed (Crick and Koch, 1998) that while this might be appropriate for animals with a limited behavioral repertoire, such as a fruit-fly, this would not do for mammals with their enormous range of behaviors. Instead evolution produced a general purpose system which can deal with all sort of contingencies and permits the animal to plan in parallel with a host of very rapid but special-purpose and stereo-typed zombie behaviors. The price is speed: consciousness reflection and planning requires a fraction of a second or longer.

4. The Nature of the Visual Representation

We argued (Crick and Koch, 1995b) that to be aware of an object or event, the brain has to construct a multilevel, explicit, symbolic interpretation of part of the visual scene. By multilevel, we mean, in psychological terms, different levels such as those that correspond, for example, to lines or eyes or faces. In neurological terms, we mean, loosely, the different levels in the visual hierarchy (Felleman and Van Essen, 1991).

The important idea is that the representation must be explicit. The word is here used in the sense that it should be easy to infer the presence or absence of the feature that is made explicit in the cellular response (one could formalize this using a perceptron).

Notice that while the information needed to represent a face is contained in the firing of the ganglion cells in the retina, the retina lacks an explicit representation for faces.

The evidence from brain damage, especially in humans, suggests that certain parts of the cortex are essential for a person to be conscious of certain aspects of the visual sensation or percept, such as color, motion, faces, etc. Zeki and Bartels (1999) have, very reasonably, described such a piece of the cortex as an essential node for that aspect of the percept, though a "key node" might be a better term. What should be considered "one aspect" must be decided by experiment. Thus "motion" is not necessarily a single aspect. Indeed there is evidence from brain damage (Vaina and Cowey, 1996) that F.D., a patient with rather limited brain damage, is impaired in the detection of second-order or texture-based motion but does perceive first-order or intensity-based motion.

The term should not be taken to imply that a person who possessed only the relevant essential node would be conscious of that aspect of the percept. It is highly probable that to produce that aspect of consciousness the node would have to interact with other parts of the brain. A node is a node, not a network. The point is that damage to that essential node would specifically remove that particular aspect of the sensation or percept, while leaving other aspects relatively intact.

The concept of an essential node is an important one. It implies that if there is no essential node for some possible aspect of consciousness the subject cannot be conscious of it. Zihl's patient (Zihl, Von Cramon and Mai, 1983) can see a car, but if the car is moving she does not see its movement because the essential nodes for movement have been damaged. So it is important not to assume that, for example, the brain can necessarily consciously detect some

particular change in its firing activity. It will not be able to do so unless there is some essential node to consciously register that type of change.

It is plausible that an explicit representation (at an essential node) might be a smallish group of neurons employing coarse coding to represent some aspect of the visual scene. In the case of a particular face, all of these neurons can fire to somewhat face-like objects (see the chapter by **Rolls**). We suggest that one set of such neurons will be all of one type, will probably be fairly close together, and will all project to roughly the same place. If all such groups of neurons (there may be several of them, stacked one above the other) were destroyed, then the person would not see a face, though he or she might be able to see the parts of a face, such as the eyes, the nose, the mouth, etc. There are other places in the brain that explicitly represent other aspects of a face, such as the emotion the face is expressing or its angle of gaze.

An attractive proposition is that the stimulus feature common to most neurons within a cortical column is what is made explicit there. Thus, orientation is one feature explicitly represented in V1, as is the direction and speed of motion in area V5. We call this the columnar hypothesis. Note that an explicit representation is a necessary but not sufficient condition for the NCC. As outlined in section 7, it matters a great deal where these neurons project.

A representation of an object or an event will usually consist of representations of many of the relevant aspects of it, and these are likely to be distributed, to some degree, over different parts of the visual system. How these properties are bound together is known as one aspect of the binding problem (see the chapter by **Singer**). It is possible that in many cases the firing of pyramidal neurons in all layers within a column that code for the object need to be synchronized for this to occur.

Much neural activity, most of which is probably unconscious, is usually needed for the brain to construct a representation. It may prove useful to consider this unconscious activity as the computations needed to find the best interpretation, while the interpretation itself may be considered to be the results of the computations, only some of which we are then conscious of. To judge from our perception, the results probably have something of a winner-take-all character.

In the past we wondered whether only some types of specific neurons might express the NCC. Silence these and a very specific loss of some specific content of consciousness might occur (like the inability to perceive a certain type of motion). A weaker version of this hypothesis holds that the NCC depends on some crucial biophysical or anatomical property of some network in the brain. Again, blocking

this feature could lead to a very specific loss of some aspect of subjective feelings.

An alternative possibility is that the NCC is necessarily global, usually identified with some emergent property of very large and diffuse assemblies of neurons. In one extreme form this would mean that, at one time or another, any neuron in cortex and associated structures could express the NCC, and it would be foolish to locate consciousness at the level of single neurons.

These possibilities are non-exclusive: it is possible that both local as well as global properties of the brain are necessary for consciousness to occur. While our thinking over the past decade has stressed particular neuronal mechanisms, that of Edelman and Tononi has focused on certain global aspects of networks (their dynamic core hypothesis; Edelman and Tononi, 2000). Both ideas might be valid.

Conscious visual representations are likely to be distributed over more than one area of the cerebral cortex and possibly over certain subcortical structures as well. We have argued that in primates---contrary to received opinion---they are not located in primary visual cortex (see below).

Whatever the NCC is, it must be present in the two hemispheres. Careful study of split brain patients demonstrates that one hemisphere usually supports language while the linguistic competence of the other is severely reduced. Yet, even this non-dominant hemisphere can clearly reason, respond purposefully and plan complex behaviors. From this one can conclude that split-brain patients have, within the confine of a single skull, two brains with two conscious minds (Sperry, 1974; Bogen, 1997).

5. What is Essential for Visual Consciousness?

The term "visual consciousness" almost certainly covers a variety of processes. When one is actually looking at a visual scene, the experience is very vivid. This should be contrasted with the much less vivid and less detailed visual images produced by trying to remember the same scene (a vivid recollection is usually called a hallucination). We are concerned here mainly with normal vivid experiences. It is possible that our dimmer visual recollections, imagery, are mainly due to the back pathways in the visual hierarchy biasing the "noise" in earlier stages of the system.

Some form of very short-term memory seems almost essential for consciousness, but this memory may be very transient, lasting for only a fraction of a second. The existence of iconic memory, as it is called, is well-established experimentally (Coltheart, 1983; Gegenfurtner and Sperling, 1993).

Psychophysical evidence for short-term memory (Potter, 1976) suggests that if we do not pay attention to some part or aspect of the visual scene, our memory of it is very transient and can be overwritten (masked) by the following visual stimulus. This probably explains many of our fleeting memories when we drive a car over a familiar route. If we do pay attention (e.g., a child crossing the road) our recollection of this can be longer lasting.

Our impression that at any moment we see all of a visual scene very clearly and in great detail is illusory, partly due to ever-present eye movements and partly due to our ability to use the scene itself as a readily available form of memory, since in most circumstances the scene usually changes rather little over a short span of time (O'Regan, 1992; see the chapter by **Simmons**). What we can grasp very rapidly and pre-attentively is the gist of a scene (Mack and Rock, 1998; Wolfe, 1998). It is this that provides us with the illusion of being able to apprehend the entire scene at once. It is likely that there are neurons at the higher visual levels whose firing directly encodes the *gist* of that scene. These neurons would then be key culprits for mediating the sensation of seeing everything at once.

Although working memory expands the time frame of consciousness, it is not obvious that it is essential for consciousness. It seems to us that working memory is a mechanism for maintaining one or a small sequence of items in consciousness. In a similar way, the episodic memory enabled by the hippocampal system (Zola-Morgan and Squire, 1993) is not essential for consciousness, though a person without it is severely handicapped.

Is attention necessary for consciousness? It seems that it is enriched by visual attention, though top-down attention is not essential for visual consciousness to occur (Rock, Linnett, Grant, and Mac, 1992; Braun and Julesz, 1998). Attention is broadly of two types: bottom-up, saliency-driven and task-independent; and top-down, produced by the planning parts of the brain. Visual attention can be directed to either a location in the visual field or to one or more objects. The neural mechanisms that achieve this are still being debated. But they are likely to involve biasing the competition among groups of cells coding for neighbouring objects in favor of the attended one (Desimone and Duncan, 1995). See our earlier discussion of neglect and Balint syndromes.

6. How does the timing of events affect the percept?

The timing of the emergence of consciousness has been called *microgenesis* (Bachmann, 2000). We can easily see something of a visual input, such as a flash of

lightening, even though it lasts for a very short time, provided it is strong enough. Bloch's law states that for stimulus duration less than about a tenth of a second (for a diffuse flash) the brightness of the stimulus appears the same provided the product of its intensity and its duration is constant. In some sense the system is integrating the input over some short time interval.

How bright does a flash of light appear if its intensity is kept the same but its duration varies from trial to trial? This can be estimated by comparing its apparent brightness with the brightness of a similar but constant light. A typical result has been described by Efron (1967). As the duration increases the light appears more intense until, for a duration of about 40 msec, it reaches a maximum after which the subjective brightness declines to a steady value (Fig. 1A). As Efron has pointed out this description can be misleading. It expresses the results of multiple trials, each one for one particular duration of flash. It does not show what a person experiences at a single trial. That is, for a flash of length of say 125 msec the subject does not see the brightness of the flash increasing rapidly and then decreasing somewhat. On the contrary she reports that she saw a steady brightness. This distinction, which has been widely overlooked, is an important one.

The results suggest that the NCC comes into being abruptly rather than gradually. Once the relevant neural activity reaches some threshold a constant percept of brightness results, at least for a short time.

Instead of a single flash of light, what do we see if a complex scene, such as people dining in a restaurant, is flashed for different short durations? The general result is that for very short exposures one perceives the general nature of the scene, the gist, as mentioned earlier. As the flash is made longer we can report more details. Once again, in any one trial we do not see the scene change. We just see more for longer flashes.

This might suggest that some of the higher levels of the visual hierarchy reach the necessary threshold for consciousness before the lower levels do. Possibly the lower levels need some feedback from the higher levels to boost their activity above threshold.

In "masking," two (or more) inputs are involved (Bachmann, 2000; Enns and DiLollo, 2000). When the two inputs blend this is sometimes referred to as *integrative masking*, though the term *blending* might be preferable. Pattern masking is when two rather different spatially superimposed patterns of contours interfere. When two patterns are not superimposed but spatially adjacent, the interference is called *metacontrast*. Masking can also occur when the masking pattern does not abut the target pattern but is nearby. This is now called object-substitution masking.

Suppose a red circular disc and an otherwise identical green circle disc are flashed simultaneously onto the same place in the retina. Not surprisingly the subject sees a yellow disc. If the red disc alone is flashed first, for only 10 msec, immediately followed by the green disc for 10 msec, the subject does not see a red disc turning green, but just a yellow disc. The yellow has a greenish tinge compared to the yellow produced when they are simultaneous. If the green disc comes first, the yellow is a little redder (Yund, Morgan and Efron, 1983). The subject perceives a mixture of the inputs, with a bias towards the later one—this turns out to be a general rule. This suggests some form of integration, with the later signal having a somewhat greater weight. If one disc appears 100 msec before the other, little blending occurs. This suggests that, in this instance, the integration time is less than 100 msec.

When masking produces interference this is because in some sense the mask is competing with the target. The subject can easily report the target if there is no mask, but if the mask is spatially superimposed on the target, and occurs about the same time, his responses are mostly at chance—he reports he did not see the target. He does better if there is a delay of 100 msec or so between the onset of the target and the onset of the mask. The mask interferes with processing of the target in the integration period leading up to consciousness. It is plausible that once some kind of neural activity due to the target has reached a certain threshold, the following mask cannot interfere with it so easily. This suggests that the conscious activity may show hysteresis---as Libet (1973) has claimed---since the activity is probably held above a threshold to some extent by some mechanism, such as loops with positive feedback (Fig. 1B). This activity correspond to the NCC for the attribute encoded by the essential node, say the target's brightness, location or color. The attribute remains consciously accessible until the activity dips below a lower threshold (Fig. 1B) at which point it becomes perceptually invisible.

In meta-contrast, the subject can see the target if the onset of the target and mask are simultaneous, but fails to see it if the onset of the mask is <u>delayed</u> by a short period, typically 50 to 100 msec or so. This is presumably because the target and the mask are initially activating different places in the brain, and it takes time for their activities to interfere with each other.

The most interesting case is object-substitution masking; that is, when the target and mask are spatially separate. No masking occurs if the target and the mask come into view simultaneously and disappear simultaneously. Masking does occur if, instead, the mask continues beyond the disappearance of the target. No

masking occurs if the target is especially salient, or if attention is focused on the target location beforehand.

The phenomenon of "masked priming" implies that (unconscious) processing of the masked target proceeds to fairly high levels. That is, some level of priming occurs whether or not the prime was masked, suggesting that the physiological correlate of priming (short-term synaptic changes?) can occur in high-level cortical areas without giving rise to consciousness.

Enns and DiLollo (2000) suggest that one is not conscious of the details of the target until the activity it has produced first reach a higher level (when there is integration) and then flows back to a lower level (possibly iteratively). They propose that it is at one of the lower levels that the activity of the sustained mask interferes with conscious seeing, presumably of the details of the pattern.

Until recently visual psychologists did not relate their results to the complex organization of the primate visual system. The study of masking in the alert monkey by neurophysiologists has only just begun (Macknik and Livingstone, 1998; Thompson and Schall, 2000). A careful study of the neurophysiological effects of masking should throw light on the processes leading to consciousness, on the integration times for the signal to reach above the thresholds for consciousness and on the ways interference works. In general the times involved seem to range from 50 to 100 msec, to 200 msec or more. This upper limit approaches the typical time between eye movements.

The picture that emerges from these speculations is a rather curious one. It bears some resemblance to Dennett's multiple drafts model (Dennett, 1991), though Dennett's ideas, though suggestive, are not precise enough to be considered scientific. The content of consciousness instead of being a continuous ever-changing flux may be more like a rapid series of 'static snapshots'. Movement, like other attributes, is signaled in a static way by special neurons. That is, movement is not symbolized by a change in time of the symbols for position, but by special neural symbols that represent movement of one sort or another. If the essential nodes for these symbols are lost, as in Zihl's patient, then though she can see the moving car she cannot see it moving.

There is little evidence for a regular clock in the brain with a mechanism that integrates over intervals of constant duration and then starts afresh over the next interval, sometime called a quantized clock. The duration of a snapshot is likely to depend on various factors, such as contrast, saliency, competition, and so on. Purves and his colleagues (Purves, Paydafar and Andrews, 1996) have described several psychological effects, such as a wagon-wheel effect under constant

illumination, which hint that there are some clock-like effects in vision. This would be worth further study.

The brain has a problem with time, since time is needed to express a symbol (by neurons firing, a process spread out over time) so it is not surprising if it symbolized changes in time, such as movements, in a special way, and differences in time in another way.

We are dealing with an intrinsically parallel system to which we have limited access introspectively. This is probably why we find it so hard to understand how it works. This does not mean that we cannot usefully analyze it into smaller parts that interact dynamically, just as the "holistic" properties of a complex organic molecule, such as a protein, can be understood by the interactions of its many amino acids and the atoms of which they are made. In the brain an essential node may be a useful unit of analysis. It may turn out that the best way to describe the NCC for any one percept is the activities at the relevant essential nodes and their dynamic interactions. This suggestion resembles the dynamic core of Edelman and Tononi (2000).

7. Bistable percepts

Perhaps the most relevant experimental approach to tracking down the footsteps of the NCC for now is to study the behavior of single neurons while the subject is looking at something that produces a bistable percept. The visual input, apart from minor eye movements, is constant; but the subject's percept can take one of two alternative forms. This happens, for example, when one looks at a drawing of the Necker cube. Which neurons follow the constant retinal input and which ones the variable percept?

A more practical alternative is to study binocular rivalry (Myerson, Miezin, and Allman, 1981). If the visual input into each eye is different, but perceptually overlapping, one usually sees the visual input as received by one eye alone, then by the other one, then by the first one, and so on. This approach has been explored by Logothetis and his colleagues by training macaque monkeys to report which of two rivalrous inputs they saw. The fairly similar distribution of switching times strongly suggests that monkeys and humans perceive these bistable visual inputs in the same way. In the first set of experiments Logothetis and Schall (1989) investigated neurons in area MT (also referred to as V5). The stimuli were vertically drifting horizontal gratings. Of the relevant neurons, only about 35% were modulated according to the monkey's reported percept. Surprisingly, half of these responded in the opposite direction to the one expected. Leopold and

Logothetis (1996) used stationary gratings. The orientation was chosen in each case to be optimal for the neuron studied, and orthogonal to it in the other eye. They recorded how the neuron fired during several alterations of the reported percept. While only a small fraction of foveal V1/V2 neurons followed the percept, about 40% of neurons in V4 did. Interestingly, about half of V1/V2 cells were anticorrelated with the stimulus.

The results of the Sheinberg and Logothetis (1997) experiments are striking. Instead of using binocular rivalry, they exploited a related paradigm: *flash suppression*. An image is presented to one eye for a second or longer. A different image is then flashed into the other eye. The former image is perceptually completely suppressed by the novel pattern, even though the first image is still present on the retina (Fig. 2). The mechanisms underlying flash suppression are unrelated to either forward masking, adaptation or any other mechanism that reduces the visibility of the ipsilateral flash (Wolfe, 1984). Flash suppression is easy to obtain, is very robust and can be exactly timed, unlike the transitions occurring during binocular rivalry.

Sheinberg and Logothetis (1997) recorded from individual neurons in the monkey's superior temporal sulcus (STS) and inferior temporal cortex (IT). Overall, approximately 90% of the relevant neurons in STS and IT reliably predict the perceptual state of the animal. That is, these neurons follow the percept rather than the retinal stimulus. Moreover, many of these neurons respond in an almost all-or-none fashion, firing strongly for one percept, yet only at background levels for the alternative one.

Although the vast bulk of electrophysiological single-cell experiments are carried out in animals, under exceptional circumstances such recordings are possible in humans. These occur in a subset of pharmacologically intractable epilepsy patients undergoing evaluation for surgical treatment. Kreiman, Koch and Fried (2000) found many neurons to be tuned for specific categories of visually presented stimuli (e.g. famous people such as politicians or entertainers, faces of unknown actors expressing various emotions, animals or cars). Many of these cells fire selectively when their preferred stimulus is perceived but not when it is perceptually suppressed during flash suppression (Kreiman, Fried and Koch, 2002). About two thirds of a total of 48 selective medial temporal lobe (MTL) neurons followed the percept (Figs. 2 and 3). No neuron responded in a significant manner to its preferred visual stimulus if this stimulus was physically present on the retina, but perceptually suppressed. That is, there is no evidence for a neuronal representation of perceptually suppressed, that is unconscious, stimuli in the

higher echelons of the visual hierarchy. Thus, both IT neurons in the macaque as well as MTL neurons in humans follow the subject's conscious percept.

Of course, without further experiments it is not possible to distinguish between correlation and causation. It does not automatically imply that just because a particular neuron follows the percept, its firing is part of the NCC. Do the MTL neurons carry a mere copy of the neuronal signal expressed already in IT for transfer into working memory or are neurons in hippocampus, amygdala and entorhinal cortex in normals necessary for conscious perception? Where are these neurons projecting to? It is, at the moment, technically difficult to do this, but it is essential to have this knowledge, or it will be almost impossible to understand the neural nature of consciousness.

8. What is the Role of V1 in Visual Consciousness

We have argued (Crick and Koch, 1995b) that one is not directly conscious of the features represented by the neural activity in primary visual cortex. Activity in V1 may be necessary for vivid and veridical visual consciousness (as is activity in the retina), but we suggest that the firing of none of the neurons in V1 directly correlates with what we consciously see.

Our reasons are that at each stage in the visual hierarchy the explicit aspects of the representation we have postulated is re-coded. We also assumed that any neurons expressing an aspect of the NCC must project directly, without re-coding, to at least some of the parts of the brain that plan voluntary action--that is what we have argued seeing is for. We think that these plans are made in some parts of frontal cortex (see below).

The neuroanatomy of the macaque monkey shows that V1 cells do not project directly to any part of frontal cortex (see the references in Crick and Koch, 1995b).

The strategy to verify or falsify this and similar hypotheses is to relate the receptive field properties of individual neurons in V1 or elsewhere to perception in a quantitative manner. Ultimately, this correlation has to be made on an individual trial-to-trial basis. If the structure of perception does not map to the receptive field properties of V1 cells, it is unlikely that these neurons directly give rise to consciousness. In the presence of a correlation between perceptual experience and the receptive field properties of V1 cells, it is unclear whether these cells just correlate with consciousness or directly give rise to it. In that case, further experiments need to be carried out to untangle the exact relationship between neurons and perception.

Psychophysical evidence is broadly consistent with our proposal. The study by He, Cavanagh and Intriligator (1996) exploit a common visual aftereffect. If a subject stares for a fraction of a minute at a horizontal grating, and is then tested with a faint grating at the same location to decide whether it is oriented vertically or horizontally, the subject's sensitivity for detecting a horizontal grating will be reduced. This adaptation is orientation specific--the sensitivity for vertical gratings is almost unchanged--and disappears quickly. He and colleagues projected a single patch of grating onto a computer screen some 25 degrees from the fixation point. It was clearly visible and their subjects showed the predictable orientation-selective adaptation effect. Adding one or more similar patches of gratings to either side of the original grating removes the lines of the grating from visibility; it is "masked." Subjectively, one still sees "something" at the location of the original grating, but one is unable to make out its orientation, even when given unlimited viewing time. Yet despite this inability to "see" the adapting stimulus, the aftereffect was as strong and as specific to the orientation of the "invisible" grating as when the grating was visible. What this shows is that visual awareness must occur at a higher stage in the visual hierarchy than orientation-specific adaptation. This aftereffect is thought to be mediated by oriented neurons in V1 and beyond, implying that the NCC must be located past this stage.

Another case in point is the observation that very high frequency gratings that are perceptually indistinguishable from a uniform field nevertheless produce robust orientation-dependent aftereffects (He and MacLeod, 2001) The presence of aftereffects from stimuli that are not consciously perceived (Blake and Fox, 1974) indicates that stimulus properties must be represented outside of awareness. Thus, activity in V1 often does not correlate with awareness.

These inferences from behavior are complemented by direct electrophysiological measurements of neuronal activity in V1 in macaque monkeys. These show that perceptual report (presumably reflecting awareness) and the mean firing activity in V1 can change completely independently. For example, disparity-selective cells in V1 differentiate between local depth cues even when these do not give rise to an overall depth percept or may respond in an identical manner to two depth cues that yield very different global depth percepts (Cumming and Parker, 1997). This suggests that single V1 cells sensitive to binocular disparity represent a first critical stage for generation of stereo cues, but that conscious perception of depth is most likely generated further upstream. A similar lack of correlation between perceptual experience and the response properties of V1 cells can be seen during blinks. They produce profound and frequent interruptions in retinal stimulation, yet barely impinge on our visual

experience. Cells in the supragranular layers of V1 show responses to visual stimulation that decrease substantially during reflex blinks (Gawne and Martin, 2000) and this reduction is significantly more pronounced than when there is an equally long gap in visual stimulation. Similarly, microsaccades that occur continually do not alter our impression of the world but are reflected in the activity of neurons in V1 (Martinez-Conde, Macknik and Hubel, 2000). Lastly, when two isoluminant colors are alternated at frequencies beyond 10 Hz, humans perceive only a single fused color with a minimal sensation of brightness flicker. In spite of the perception of color fusion, color opponent cells in primary visual cortex of two alert macaque monkeys follow high-frequency flicker well above heterochromatic fusion frequencies (Gur and Snodderly, 1997). In other words, neuronal activity in V1 can clearly represent certain retinal stimulation that is not perceived.

Changes in conscious perception can also occur in the absence of concomitant variations of neuronal activity in V1. As mentioned above, the majority of the V1/V2 cells recorded by Leopold and Logothetis (1996) do not co-vary with the animal's behavior during binocular rivalry. Instead, their firing reflects stimulus properties. Recently, two neuroimaging studies in humans have presented evidence arguing for a stronger role for V1 in binocular rivalry and hence, by implication, visual awareness. The most popular form of functional neuroimaging is Blood Oxygenation-Level-Dependent (BOLD) functional MRI. BOLD measures a component of the hemodynamic response associated with local neural activity (for the exact relationship between the two, see Logothetis, Pauls, Augath, Trinath and Oeltermann, 2001). Averaged over the whole of human V1, reliable fluctuations in BOLD contrast activity are observed (Polonsky, Blake, Braun and Heeger, 2000). The amplitude of these fluctuations is about half as large as those evoked by non-rivalrous stimuli, and even larger in the part of V1 associated with the blind spot representation which contains cells receiving only monocular input from the non-blind spot eye (Tong and Engel, 2001). The reasons for this discrepancy are not yet clear but might well involve the complex nature of the relationship between single cell activity and BOLD measurements (Logothetis, Pauls, Augath, Trinath and Oeltermann, 2001). In particular, it is possible that feedback activity from higher cortical areas to lower ones does give rise to metabolic demand in the lower area due to the associated synaptic activity and so results in enhanced BOLD contrast. Yet the feedback is, by itself, not sufficiently powerful to give rise to spiking in the V1 cells. Other fMRI BOLD studies in humans show a poor correlation between V1 activity and visual awareness (for a review, see Rees, Kreiman and Koch, 2002).

Note that our hypothesis is not a simple one to prove, since it applies to all V1 cell populations. Furthermore, the existence of a strong relationship between V1

activity and percept does not, of course, imply causation. We hope that further neuroanatomical work will make our hypothesis plausible for humans. If correct, it would narrow the search to areas of the brain farther removed from the sensory periphery.

9. The Frontal Lobe Hypothesis

As mentioned several times, we hypothesize that the NCC must have access to explicitly encoded visual information and directly project into the planning stages of the brain, associated with the frontal lobes in general and with prefrontal cortex in particular (Fuster, 1997). We would therefore predict that patients unfortunate enough to have lost their entire prefrontal cortex on both sides (including Broca's area) would not be visually conscious, although they might still have well-preserved, but unconscious, visual-motor abilities. No such patient is known to us (not even Brickner's famous patient; for an extensive discussion, see Damasio and Anderson, 1993). The visual abilities of any “frontal lobe” patient needs to be carefully evaluated using a battery of appropriate psychophysical tests.

The fMRI study of the blindsight patient G.Y. (Sahraie *et al.*, 1997) provides direct evidence for our view by revealing that prefrontal areas 46 and 47 are active when G.Y. is aware of a moving stimulus.

Findings of neurons in the inferior prefrontal cortex (IPC) of the macaque that respond selectively to faces--and that receive direct input from regions around STS and the inferior temporal gyrus that are well known to contain face-selective neurons--is very encouraging (Scalaidhe, Wilson and Goldman-Rakic, 1997). This raises the question of why would face cells be represented in both IT and IPC. It is unlikely that exactly the same information is represented at two different locations so it will be important to find out in what way they differ.

Large-scale lesion experiments carried out in the monkey suggest that the absence of frontal lobes leads to complete blindness (Gazzaniga, 1966; Nakamura and Mishkin, 1986). One would hope that future monkey experiments reversibly (by cooling and/or injection of GABA agonists) inactivate specific prefrontal areas and demonstrate the specific loss of abilities linked to visual perception while visual-motor behaviors -- mediated by the on-line system -- remain intact.

Some cognitive scientists (in particular Jackendoff, 1987) have argued for an “Intermediate-Level” theory of consciousness in which neither early sensory processing (the "outer world") nor high-level three-dimensional information or thoughts (the "inner world") are directly accessible to consciousness. The

Intermediate-Level" hypothesis raises the intriguing possibility that although much of the frontal lobes might be necessary for consciousness, the NCC might be preferentially located in more intermediate levels of cortex, such as the inferior temporal lobes (for a more detailed discussion of this and some experimental ramifications, see Crick and Koch, 2000).

10. Where are We and Where Should We be Going?

We reviewed the evidence in favor of the hypothesis that there might be anatomically localized neurons or neural pathways whose level of spiking activity mediates consciousness. The NCC is not likely to be found in V1, but among neurons in the ventral pathway (in particular in and around area IT), and in its recipient zones in the medial temporal and frontal lobes. Neurons from posterior parietal areas can modulate the competition in this stream, without being strictly necessary for conscious perception.

It is likely that a conscious percept requires that the relevant activity in the relevant neurons exceeds some minimal amount. It is possible that such a threshold can only be exceeded by some type of feedback activity from a higher cortical area (Pollen, 1999; Lamme and Roelfsema, 2000). Furthermore, this minimal activity might have to persist for a minimum amount of time.

Yet the average firing activity need not be the only relevant variable. An alternate hypothesis is that consciousness might be associated with specific types of neural activity, such as oscillatory or synchronized discharges. These two hypotheses are not mutually exclusive, and a combination of synchronous activity and neural spike rates might be an important attribute of the NCC. Indeed, we suggested early on that synchronized oscillations in the broad "40 Hz" range may underlie feature integration and form a substrate for visual consciousness, although the idea remains controversial (Crick and Koch, 1990; Engel and Singer, 2001; see the chapter by **Singer**). In general, a synchronized neuronal population will be more efficient in evoking a strong postsynaptic response compared to a network of independently firing neurons; that is, it carries a more powerful postsynaptic 'punch'. The brain may use correlated firing for more than one purpose. In particular, it might embody the effects of top-down attention. As consciousness frequently involves the suppression of competing percepts, such synchronized and oscillatory firing might often accompany the NCC without being strictly necessary. Conversely, synchronized firing might also underlie other processes, such a figure-ground segregation, so that their mere presence might not reliably indicate a NCC.

At the moment, the most promising experiments to discover the correlates of visual consciousness are those on bistable percepts which need to be expanded to include other brain areas (in particular prefrontal cortex) in both the monkey and in humans. It is important to discover which neurons express the NCC in each case (which neuronal subtype, in what layer, and so on), how they fire (do they fire in bursts; are discharge patterns across neurons synchronized) and, especially, where they project to. To assist this, more detailed neuroanatomy of the connectivity will be needed. It is also important to discover how the various on-line systems work, so that one can contrast their (unconscious) neuronal activity with the NCC (Koch and Crick, 2001).

It will not be enough to show that certain neurons embody the NCC in certain--limited--visual situations. Rather, we need to identify the NCC for a sufficiently large and representative sample of visual inputs. One popular approach is to study the many visual illusions. For instance, humans clearly perceive, under appropriate circumstances, a transient motion aftereffect. On the basis of fMRI it has been found that the human equivalent of cortical area MT is activated by the motion aftereffect (in the absence of any moving stimuli; Tootell *et al.*, 1995). The time course of this illusion parallels the time course of activity as assayed using fMRI. In order to really pinpoint the NCC, one would need to identify individual cells expressing this, and similar, visual aftereffects.

This is not to argue that studying visual consciousness constitutes the one and only means of cracking the mind-brain problem. Investigating other forms of consciousness might have practical advantages (for instance, the awareness of the exact US-CS stimulus contingency during certain forms of associative conditioning; Clark and Squire, 1998). In order to understand consciousness, we need to understand cerebral cortex. We can only hope to achieve this goal by deliberately interfering with cortex and its components in a delicate, reversible and transient manner. Molecular biology is beginning to provide the relevant tools (e.g. Lois, Hong, Pease, Brown, and Baltimore, 2002). Whether these are best applied to the macaque's visual system or to another modality in the more accessible mouse remains an open question.

Finding the NCC would only be the first, albeit critical, step in understanding consciousness. The next biggest challenge would be to bridge the gap between correlation and causation. For this to occur, we would need to understand the postsynaptic action of neurons expressing the NCC, what happens to them in various diseases known to affect consciousness, such as schizophrenia or autism, and so on. And, of course, a final theory of consciousness would have to explain the central mystery: how a physical system with a particular

architecture can have feelings and qualia (Chalmers, 1995). After several thousand years of speculation, it would be very gratifying to solve this puzzle.

Acknowledgements: We thank the J.W. Kieckhefer Foundation, the W.M. Keck Foundation Fund for Discovery in Basic Medical Research at Caltech, NIH, NIMH, and NSF for support.

References

Bachmann, T. 2000. Microgenetic approach to the conscious mind, John Benjamins Publishing Company: Amsterdam and Philadelphia.

Baron-Cohen, S. 2001. Consciousness of the physical and the mental. In: Finding Consciousness in the Brain, P.G. Grossenbacher, ed., pp. 61-76. John Benjamins: Amsterdam.

Blake, R. and R. Fox, 1974. Adaptation to invisible gratings and the site of binocular rivalry suppression, Nature, 249: 488-490.

*Bogen, J. E. 1997. Some neurophysiologic aspects of consciousness, Seminar Neurobiol., 17: 95-103.

Boussaoud, D., G. Di Pellegrino, and S. P. Wise, 1996. Frontal lobe mechanisms subserving vision-for-action versus vision-for-perception, Behav. Brain Res., 72: 1-15.

Braun, J., and B. Julesz, 1998. Withdrawing attention at little cost: Detection and discrimination tasks, Percep. Psychophys., 60: 1-23.

Chalmers, D., 1995. The Conscious Mind: In Search of a Fundamental Theory, Oxford University Press: Oxford, United Kingdom.

Clark, R. E., and L.R. Squire, 1998. Classical conditioning and brain systems: the role of awareness, Science, 280: 77-81.

Coltheart, M., 1983. Iconic memory, Phil. Trans. R. Soc. Lond. B, 302: 283-294.

Crick, F., and C. Koch, 1990. Towards a neurobiological theory of consciousness, Seminars Neuroscience, 2:263-275.

Crick, F., and C. Koch, 1995a. Why neuroscience may be able to explain consciousness, Scientific American. 273: 84-85.

Crick, F., and C. Koch, 1995b. Are we aware of neural activity in primary visualcortex? Nature, 375: 121-123.

*Crick, F., and C. Koch, 1998. Consciousness and neuroscience, Cerebral Cortex, 8: 97-107.

Crick, F., and C. Koch, 2000. The unconscious homunculus. With commentaries by multiple authors, Neuro-Psychoanalysis, 2: 3-59.

Cumming, B. G., and A. J. Parker, 1997. Responses of primary visual cortical neurons to binocular disparity without depth perception, Nature, 389: 280-283.

Damasio, A. R., and S. W. Anderson, 1993. The frontal lobes. In: Clinical Neuropsychology, 3rd ed. K. M. Heilman and E. Valenstein, eds. pp. 409-460. Oxford University Press: Oxford, United Kingdom.

*Dennett, D.C. Consciousness Explained. Little, Brown and Company: Boston, Massachusetts.

*Desimone, R., and J. Duncan, 1995. Neural mechanisms of selective visual attention, Ann. Rev. Neurosci., 18: 193-222.

Distler, C., D. Boussaoud, R. Desimone, and L. G. Ungerleider, 1993. Cortical connections of inferior temporal area TEO in macaque monkeys, J. Comp. Neurol., 334:125-150.

*Driver, J., and J.B. Mattingley, 1998. Parietal neglect and visual awareness, Nature, Neuroscience, 1: 17-22.

*Edelman, G.M. and G. Tononi, 2000. A Universe of Consciousness, Basic Books: New York, New York.

Efron, R., 1967. The Duration of the Present. In: Interdisciplinary Perspectives of Time. Editors, pp. 713-729. Annals of the New York Academy of Sciences: New York.

*Engel, A. K. and W. Singer, 2001. Temporal binding and the neural correlates of sensory awareness, Trends Cognitive Science, 5: 16-25.

*Enns, J.T., and V. DiLollo, 2000. What's new in visual masking?, Trends Cognitive Science 4**:** 345-352.

*Felleman, D. J., and D. Van Essen, 1991. Distributed hierarchical processing in the primate cerebral cotrex, Cerebral Cortex, 1: 1-47.

Franz, V.H. et al. 2000. Grasping visual illusions: No evidence for a dissociation between perception and action, Psychol. Sci., 11: 20-25.

Fuster, J. M., 1997. The Prefrontal Cortex: Anatomy, Physiology, and Neuropsychology of the Frontal Lobe, 3rd ed. Philadelphia: Lippincott-Raven.

Gazzaniga, M. S., 1966. Visuomotor integration in split-brain monkeys with other cerebral lesions, Exp. Neurol., 16: 289-298.

Gegenfurtner, K.R., and G. Sperling, 1993. Information transfer in iconic memory experiments, J. Exp. Psych.: Human Percep. Perform., 19: 845-866.

Gur, M., and D. M. Snodderly, 1997. A dissociation between brain activity and perception: chromatically opponent cortical neurons signal chromatic flicker that is not perceived. Vision Res., 37:377-382.

He, S., P. Cavanagh, and J. Intriligator, 1996. Attentional resolution and the locus of visual awareness. Nature, 383:334-337.

He, S. and D.R. Macleod, 2001. Orientation-selective adaptation and tilt after-effect from invisible patterns, Nature, 411: 473-476.

*Jackendoff, R. 1987. Consciousness and the Computational Mind, MIT Press: Cambridge, Massachusetts.

Karnath, H.O. 2001. New insights into the functions of the superior temporal cortex, Nature Reviews Neuroscience, 2: 568-576.

Koch, C., and F.C. Crick, 2000. Some Thoughts on Consciousness and Neuroscience. In: The Cognitive Neurosciences, 2-nd edition, Gazzaniga, M.S., editor. pp. 1285-1294. MIT Press: Cambridge, Massachusetts.

Koch, C. and F. Crick, 2001. On the zombie within, Nature, 411: 893.

Kreiman, G., I. Fried, and C. Koch, 2002. Single neuron correlates of subjective vision in the human medial temporal lobe, Proc. Natl. Acad. Sci. USA, in press.

Kreiman, G., C. Koch, and I. Fried, 2000. Category-specific visual responses of single neurons in the human medial temporal lobe, Nature Neuroscience, 3: 946-953.

Lamme, V. A. and P.R. Roelfsema, 2000. Two distinct modes of vision offered by feedforward and recurrent processing. Trends Neuroscience, 23: 571-579.

Leopold, D. A., and N. K. Logothetis, 1996. Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry, Nature, 379: 549-553.

Libet, B. 1973. Electrical stimulation of cortex in human subjects and conscious sensory aspects. In: Handbook of Sensory Physiology, Vol. II: Somatosensory Systems, Iggo, A., Editors, pp. 743-790. Springer Verlag: Berlin, Germany.

*Logothetis, N. K., J. Pauls, M. Augath, T. Trinath, and A. Oeltermann, 2001. Neurophysiological investigation of the basis of the fMRI signal, Nature, 412: 150-157,

Logothetis, N., and J. Schall, 1989. Neuronal correlates of subjective visual perception, Science, 245: 761-763.

Lois, C., E.J. Hong, S. Pease, E.J. Brown, and D. Baltimore 2002. Germline transmission and tissue-specific expression of transgenes delivered by lentiviral vectors, Science, 295: 868-872.

Mack A. and I. Rock, 1998. Inattentional Blindness. MIT Press: Cambridge, Massachusetts.

Macknik, S.L., and M.S. Livingstone, 1998. Neuronal correlates of visibility and invisibility in the primate visual system, Nature Neuroscience 1**:** 144-149.

*Milner, D., and M. Goodale, 1995. The Visual Brain in Action. Oxford University Press: Oxford, United Kingdom.

Myerson, J., F. Miezin, and J. Allman, 1981. Binocular rivalry in macaque monkeys and humans: a comparative study in perception, Behav. Anal. Lett., 1: 149-156.

Nakamura, R. K., and M. Mishkin, 1986. Chronic blindness following lesions of nonvisual cortex in the monkey. Exp. Brain Res., 62:173-184.

Nimchinsky, E.A., E. Gilissen, J.M. Allman, D.P. Perl, J.M. Erwin, and P.R. Hof, 1999. A neuronal morphologic type unique to humans and great apes, Proc. Natl. Acad. Sci. USA, 96: 5268-5273.

O'Regan, J. K., 1992. Solving the "real" mysteries of visual perception: The world as an outside memory, Canadian J. Psychol., 46: 461-488.

Penfield, W. and H. Jasper, 1954. Epilepsy and the Functional Anatomy of the Human Brain. Little, Brown and Company: Boston, Massachusetts.

Pollen, D. A. 1999. On the neural correlates of visual perception. Cerebral Cortex, 9: 4-19.

Polonsky, A., R. Blake, J. Braun J, and D.J. Heeger 2000. Neuronal activity in human primary visual cortex correlates with perception during binocular rivalry, Nature Neuroscience, 3: 1153-1159.

Potter, M. C., 1976. Short-term conceptual memory for pictures, Exp. Psychol.: Hum. Learn. Mem., 2: 509-522.

Purves, D., J.A., Paydartar, and T.J. Andrews, 1996. The wagon wheel illusion in movies and reality, Proc. Natl. Acad. Sci. USA, 93: 3693-3697.

Rafal, R.D. 1997. Balint syndrome. In: Behavioral Neurology and Neuropsychology, Feinberg, T.E. and M.J. Farah, eds., pp. 337-356. McGraw-Hill: New York, New York.

*Rees, G., G. Kreiman, and C. Koch, 2002. Neural correlates of consciousness in humans. Nature Reviews Neuroscience, 3: 261-270.

Revonsuo, A., M. Johanson, J.-E. Wedlund, and J. Chaplin, 2000. The zombie among us. In: Beyond Dissociation, Rossetti, Y., and A. Revonsuo, eds., pp. 331-351. John Benjamins: Amsterdam, Netherlands.

Rock, I., C. M. Linnett, P. Grant, and A. Mack, 1992. Perception without attention: Results of a new method, Cogn. Psychol., 24: 502-534.

Rossetti, Y. 1998. Implicit short-lived motor representations of space in brain damaged and healthy subjects, Consciousness and Cognition, 7: 520-558.

Sahraie, A., L. Weiskrantz, J.L. Barbur, A. Simmons, S.C.R. Williams, and M.J. Brammer, 1997. Pattern of neuronal activity associated with conscious and unconscious processing of visual signals, Proc. Nat. Acad. Sci. USA, 94: 9406-9411.

Scalaidhe, S.P.O., F.A.W. Wilson, and P.S. Goldman-Rakic, 1997. Areal segregation of face-processing neurons in prefrontal cortex, Science, 278:1135-1138.

Schacter, D.L. 1989. On the relation between memory and consciousness: Dissociable interactions and conscious experience. In: Varieties of Memory and Consciousness: Essays in Honor of Endel Tulving. Roediger, H.L., and F.I.M. Craik, eds., pp. 355-389. Lawrence Erlbaum Associates: Hillsdale, New Jersey.

Schein, S. J., and R. Desimone, 1990. Spectral properties of V4 neurons in the macaque, J. Neuroscience, 10: 3369-3389.

Sheinberg, D. L., and N. K. Logothetis, 1997. The role of temporal cortical areas in perceptual organization, Proc. Natl. Acad. Sci. USA, 94: 3408-3413.

Sperry, R.W. 1974. Lateral specialization in the surgically separated hemispheres. In: Neuroscience 3-rd Study Program, Schmitt, F.O., and Worden, F.G., eds., MIT Press, Massachusettes, Cambridge.

Thompson, K.G., and J.D. Schall, 2000. Antecedents and correlates of visual detection and awareness in macaque prefrontal cortex, Vision Res., 40: 1523-38.

Tong, F., and S.A. Engel, 2001. Interocular rivalry revealed in the human cortical blind-spot representation, Nature, 411: 195-199.

Tootell, R. B. H., J. B. Reppas, A. M. Dale, R. B. Look, M. I. Sereno, R. Malach, T. J. Brady, and B. R. Rosen, 1995. Visual motion aftereffect in human cortical area MT revealed by functional magnetic resonance imaging, Nature, 375: 139-141.

*Ungerleider, L. G., and M. Mishkin, 1982. Two cortical visual systems. In: Analysis of Visual Behavior. D. J. Ingle, M. A. Goodale, and R. J. W. Mansfield, eds. pp. 549-586. MIT Press: Cambridge, Massachusetts.

Vaina, L., and A. Cowey, 1996. Impairment of the perception of second order motion but not first order motion in a patient with unilateral focal brain damage, Proc. Royal Society London B, 263: 1225-1232.

*Velmans, M. 1991. Is human information processing conscious?, Behav. Brain Sciences, 14: 651-726.

Volkmann, F. C., L. A. Riggs, and R. K. Moore, 1980. Eye-blinks and visual suppression, Science, 207:900-902.

Webster, M. J., J. Bachevalier, and L. G. Ungerleider, 1994. Connections of inferiortemporal areas TEO and TE with parietal and frontal cortex in macaque monkeys, Cerebral Cortex, 5: 470-483.

Westheimer G. 1999. Gestalt theory reconfigured: Max Wertheimer's anticipation of recent developments in visual neuroscience, Perception, 28: 5-15.

Wolfe, J., 1984. Reversing ocular dominance and suppression in a single flash, Vision Res., 24: 471-478.

Wolfe, J. M., 1998. Visual Memory: What do you know about what you saw?, Current Biology, 8(9): R303-R304.

Yund, E. W., H. Morgan, and R. Efron, 1983. The micropattern effect and visible persistence, Perception & Psychophysics, 34(3): 209-213.

Zeki, S., and A. Bartels, 1999. Toward a theory of visual consciousness, Consciousness and Cognition, 8: 225-259.

Zihl, J., D. Von Cramon, and N. Mai, 1983. Selective disturbance of movement after bilateral brain damage, Brain, 106:313-340.

Zola-Morgan, S., and L. R. Squire, 1993. Neuroanatomy of memory, Ann. Rev. Neuroscience, 16: 547-563.

Legends

Figure 1: How does a conscious percept and the neural activity that encodes it evolve in time? **(A)** Schematic drawing of the relationship between subjectively perceived brightness and stimulus duration. A flash of intermediate duration is perceived as brighter as a flash of much longer duration. Yet the percept associated with a single trial does not wax and wane as suggested by this function; instead it remains constant. **(B)** Hypothetical time course of the critical activity at the essential node coding for the brightness percept. Once this exceeds some amplitude, this NCC activity is sustained enough to project widely throughout cortex and the subject becomes consciousness of the stimulus and its brightness. Awareness is expressed until the NCC drops below some level (the two threshold values don't have to be identical, the phenomena of hysteresis).

Figure 2: Example of a response of a single neuron in the left parahippocampal gyrus of a human patient during flash suppression. **(a)** A photo of one of the Beatles, Paul McCartney, is projected into the right eye of the patient. After one second, a picture of a house is flashed into the left eye. This novel image completely suppresses the percept of McCartney and only the house is consciously seen. **(b)** The converse situation: a horizontal grating is shown to the left eye. After 1 sec, the image of McCartney is flashed into the right eye, completely suppressing the grating percept, although this image is still present on the left retina. The particular neuron recorded here follows the conscious percept of the patient, rather than the retinal stimulus. It only fires to its preferred stimulus----McCartney---whenever the subject perceives this stimulus. From Kreiman, Fried and Koch (2002).

Figure 3: Averaged normalized spike density functions of twelve medial temporal lobe (MTL) neurons in human patients during flash suppression that followed the percept. These cells responded selectively to a specific face or image in the test set (as in Fig. 2). The light trace corresponds to the response for all trials when a stimulus from the cell's effective category was shown monocularly for 1 sec **(a)** or during the 0.5 sec long flash presentation period **(b)**. The dark trace corresponds to the response to all other, ineffective, stimuli (shaded regions are the 95% confidence intervals). As all subjects perceive the effective stimuli during the monocular presentation (in **a**) or during flash suppression (in **b**), the cellular responses closely follow the subject's conscious visual experience. From Kreiman *et al* (2002).

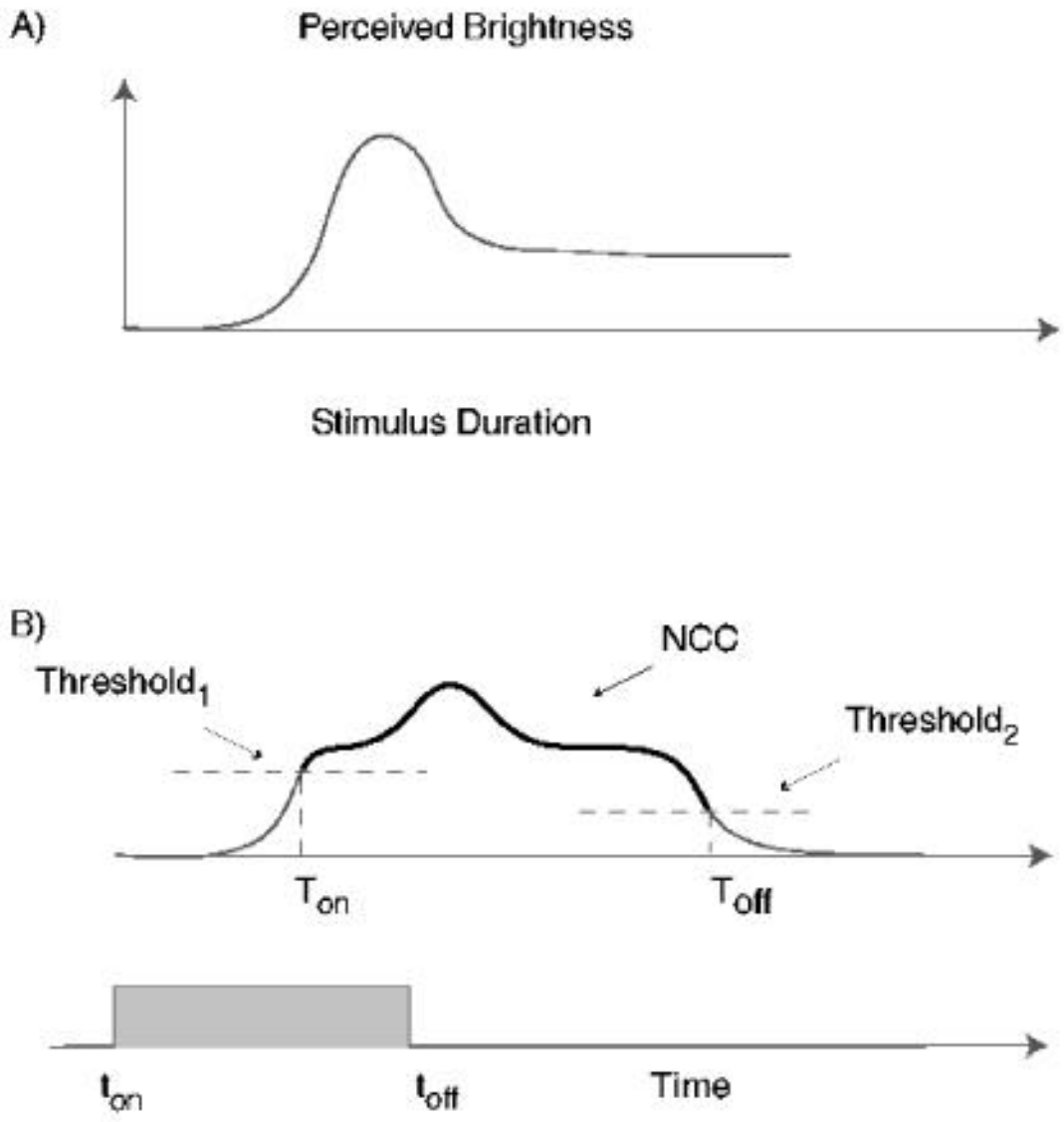
A)
Perceived Brightness
Stimulus Duration
B)
NCC
Threshold1
Threshold2
Ton
Toff
ton
toff
Time

FIGURE 1

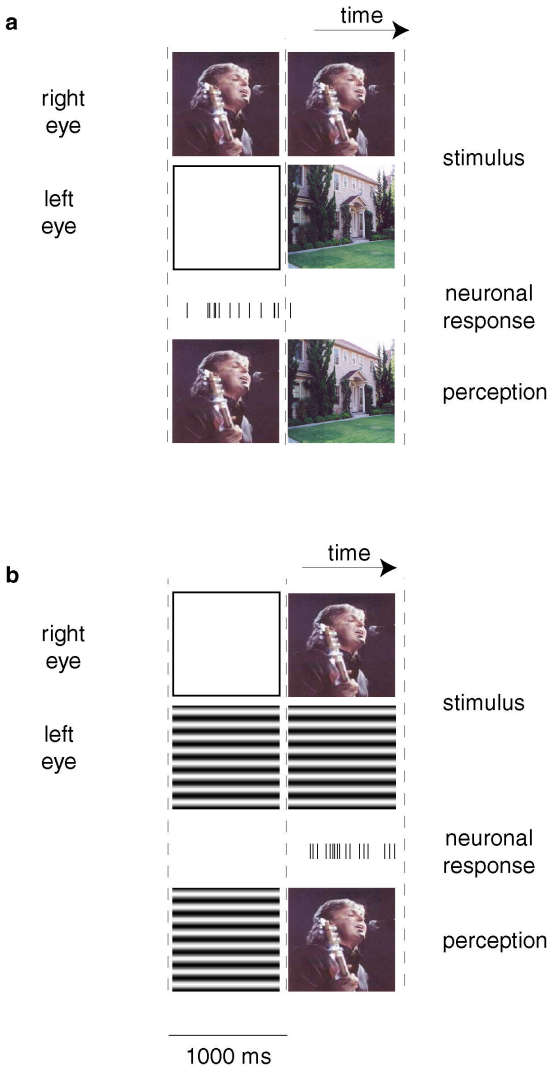
a
time
right eye
left eye
stimulus
neuronal response
perception
b
time
right eye
left eye
stimulus
neuronal response
perception
1000 ms

FIGURE 2

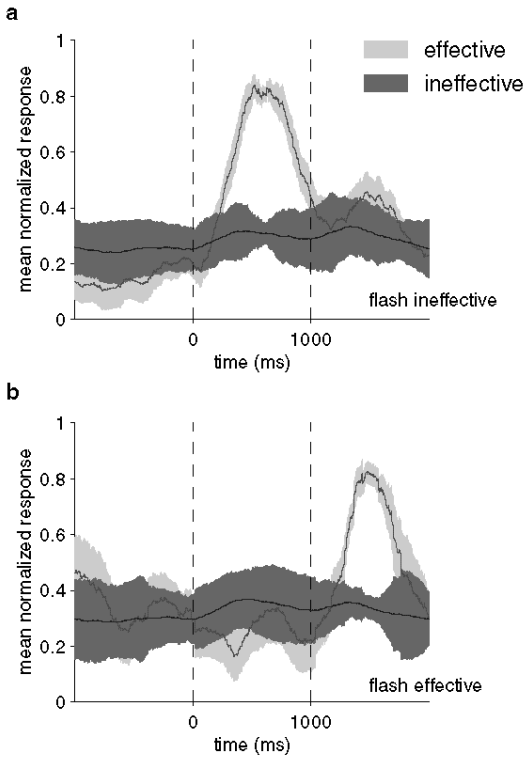
a
effective
ineffective
mean normalized response
1
0.8
0.6
0.4
0.2
0
0
1000
time (ms)
flash ineffective
b
mean normalized response
1
0.8
0.6
0.4
0.2
0
0
1000
time (ms)
flash effective

FIGURE 3